Capstone Project Submission

Instructions:
i) Please fill in all the required information.
ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:

1.Naman Veer(ravashing.naman@gmail.com)
- Data Cleaning
- Handling the null values
- Brainstorming and leading the team.

2.Pulukuri Neil Samuel Sadanand (neilsamuel509@gmail.com)
- Analysis on Price of rooms
- Analysis on room types
- Analysis on Availability of rooms

3.Nikhileshwar Karamala
- Correlation
- Analysis on Hosts

4.Sarvana Adithya (sarvanaadithya125@gmail.com)
- Analysis on Outliers
- Visualization

Please paste the GitHub Repo link.

Github Link:- https://github.com/neilsamuel123/pract_alma_better.git

---

Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)

Project Summary

The project is Eda on AIR BNB data set. AIR BnB is a company which deals with lodging facilities starting from renting shared room to renting an apartment. The following steps were performed.

1. **First step is to import the data to working environment. So pandas were used.**

2. **Secondly the NaN values had to be removed. The column "number_of_reviews" had around 10000 null values and this column was used in data analysis. So the NaN values were replace by 0.**

3. **Thirdly the analysis was done across the five neighborhood groups. The following are the conclusions**
   1. **Manhattan had the greatest number of properties.**
   2. **Manhattan was the most expensive area.**
   3. **Home/Apartments were the most expensive room types.**
   4. **Manhattan had more Home/Apartments, Brooklyn and Queens were dominated by private rooms while Staten Island and Bronx Island had equal proportions of Private rooms and Shared rooms because these are Islands.**
   5. **Total Price = minimum nights to stay in room * price**

6. Sonder NYC has the greatest number of properties across the 5 neighborhood groups yet she isn't seen as a likeable host because she rented Home/Apartments mostly which were on the monthly rents (29 days).
7. Dona was the most likeable hosts across the neighborhood groups who owned only two private rooms in total yet most likeable because the room was 47$ and the room was close to New York.
8. Price isn't related with reviews but availability of rooms is correlated to reviews.
9. Manhattan is the group which reviewed the most.

4. Now the analysis was done for each neighborhood group.
   1. **Brooklyn Neighborhood Group**
      1. The most likeable hosts was Asa because she rented Home apartment for 160$ and the apartment was available for 269 days throughout the year.
   2. **Manhattan Neighborhood Group.**
      1. The most likeable hosts was Jj because he rented private room in Harlem District for 49$ and the room was available for 339 days throughout the year.
   3. **Queens Neighborhood Group**
      1. The most likeable hosts was Dona because she rented private room for 47$ and the room was available for 333 days throughout the year.
   4. **Staten Island Neighborhood Group**
      1. The most likeable hosts was Lissette because she rented private room for 36$ and the room was available for 340 days throughout the year.
   5. **Bronx Island Neighborhood Group.**
      1. The most likeable hosts was Chandra because she rented private room for 53$ and the room was available for 43 days throughout the year